

DOCUMENT RESUME

ED 268 872

HE 019 237

AUTHOR Tompkins, Loren D.; Mehring, Teresa
TITLE Those Standardized Competency Examinations: What Do They Really Tell Us about Our Student Bodies? ASHE 1986 Annual Meeting Paper.
PUB DATE Feb 86
NOTE 30p.; Paper presented at the Annual Meeting of the Association for the Study of Higher Education (San Antonio, TX, February 20-23, 1986).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Ability; *Achievement Tests; *Competence; Educational Testing; Higher Education; Language Tests; Mathematics Tests; *Outcomes of Education; Reading Tests; Student Characteristics; *Student Evaluation; Testing; *Undergraduate Students
IDENTIFIERS *ASHE Annual Meeting; Kansas; *Nelson Denny Reading Tests; Preprofessional Skills Tests

ABSTRACT

Scholastic and personal characteristics of students undertaking exit competency examinations are investigated. Research questions concerned: what the tests measure, whether the tests provide unique information or are duplicating other easily obtained measures, whether it is possible to increase the probability of student success by controlling early higher educational experiences, and whether a workable definition of competency can be inferred. Analysis was undertaken for 1,195 Kansas public university undergraduates who took the Pre-Professional Skills Tests in Math and English and the Nelson-Denny Reading Achievement Test. Information from transcripts was obtained on American College Testing Program scores, general education enrollment patterns, student classification, grade point average, and major. Based on multiple regression analysis, it was found that competency as defined by tests is a pure function of scholastic ability only at the extremes of the distribution, with personal factors manifesting increased importance in the definition of competency for the majority of the students involved. Included are suggested actions that can be taken by faculty, student service providers, and administrators in dealing with public demands for quality assurance. (SW)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED268872

Those Standardized Competency Examinations:
What Do They Really Tell Us About Our Student Bodies?

Loren D. Tompkins
Emporia State University

Teresa Mehring
Emporia State University

February, 1986

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

- Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Loren D. Tompkins

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

HE 019 237



Association for the Study of Higher Education

The George Washington University/One Dupont Circle, Suite 630/Washington, D.C. 20036
(202) 296-2597

This paper was presented at the Annual Meeting of the Association for the Study of Higher Education held at the Gunter Hotel in San Antonio, Texas, February 20-23, 1986. This paper was reviewed by ASHE and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC collection of ASHE conference papers.

Abstract

This study conducted at a regional midwestern university involved the consideration of transcripts, personal data, and standardized competency test scores of 1195 underclassmen. Variables relating to scholastic aptitude and personal characteristics were regressed on competency test scores. Results indicated that competency as defined by tests is a pure function of scholastic ability only at the extremes of the distribution, with personal factors manifesting increased importance in the definition of competency for the majority of the students involved. The study suggests the relative importance of the factors studied and considers actions which can be taken by faculty, student service providers, and administrators in dealing with public demands for quality assurance.

A burning issue in higher education in the mid 19th century was the place of the sciences and practical arts in the curriculum. Traditionalists bemoaned the retreat from the classical program as a weakening of the intellectual dimensions of the university. While a great deal of change and reform has occurred since that time, the basic premise of discussion, what constitutes a quality higher education program, has not changed.

The Truman Commission Report in 1947 as well as the inception of government subsidized veteran education set the stage for a movement toward a quasi-universal philosophy of higher education. Cross (1971) points to the explosive growth of higher education as the nation's colleges increased enrollment from approximately 3% of the college-aged population in 1937 to 37% in 1970. She also cites the change in student body composition from a traditional male, white, Anglo-Saxon, Protestant base, whose roots lay in the higher socioeconomic strata of society to a more equalitarian mix which includes women, minorities, low achievers and the impoverished.

At the same time that this general opening of academe was occurring societal pressures were also being exerted upon the very nature of the institution. The 1960's brought a strong movement toward student involvement in governance, similar in many ways to the Populist Movement in government. Curriculum was impacted in many ways, the most notable being the initiation of the non-traditional majors and BGS degrees. At

COMPETENCY TESTING

approximately the same time three other issues emerged--a heightened public awareness of, and involvement with, the university itself (Magarrel, 1982; Biemiller, 1982; Jacobsen, 1985); a declining enrollment base (Mingle, 1981); and rapidly declining college entrance examination scores (National Commission on Excellence in Education, 1983). The outcome of all of these factors has been a general questioning of the quality of the product being certified by institutions of higher education in general, as well as public demands for quality assurance.

The institution itself has been aware of the inherent quality problems for years. Fritz Machlup sounded a warning cry concerning falling quality standards as early as 1970, and his concerns were soon echoed by others (Bestor, 1974). Meanwhile, attempts were being made to insure quality on a grand scale through complete reorganization in some states (Moos, 1981). However, another very different approach has been taken in other areas, often due to external political pressures. A major thrust in this regard has been the practice of manipulating numerical or quantifiable standards to assess both individual student and program quality. In various settings GPA requirements, admissions test cut-scores and credit hour totals have been used, but a common approach has been to adopt norm referenced standardized test for purposes of assuring quality, or establishing proficiency or competency, depending upon the local situation. This approach

COMPETENCY TESTING

to quality assessment has long been associated with public elementary-secondary education. Thus, it is very familiar, and, on first glance, appealing for its simplicity. As Astin (1982) points out, the use of a standardized test as a quality measure leads to ranking of institutions with the implicit assurance that the higher the average test scores, the higher the quality of the institution. While some type of mandated proficiency testing existed in one quarter of the states in 1980 (Kraetsch, 1980), such requirements were in place in over half by 1983 (Watkins, 1983). That such usage of tests is questionable is indicated by the Educational Testing Service's warning that the practice of using standardized test results to judge competency is not justified unless such usage is independently validated by comparison to known criteria (Preer, 1983).

This study is a statistical investigation of the scholastic and personal characteristics of a large group of students involved in an exit competency examination program. Instruments used were the Pre-Professional Skills Tests in Math and English, and the Nelson-Denny Reading Achievement Test. The initial research was designed to validate artificially developed cut-off scores, but it soon became apparent that such an approach was simplistic. The thrust was then changed to study a much more basic set of research questions:

1. What are these tests measuring, or in other words, what are the tests suggesting about our students?

COMPETENCY TESTING

2. Are the tests generating unique information or are they redundant with other easily obtained measures?
3. Is it possible to increase the probability of student success by more carefully controlling his/her early higher educational experiences?
4. By looking at the interrelationships between various quality indicators and the tests, can a workable definition of competency be inferred?

Data Base

The university setting in which the data was collected requires every student to achieve a "passing" score on the PPST Math and Writing Examinations, and on the Nelson-Denny Reading Achievement Test, as a condition of graduation. The cut scores are defined as 172 on the PPST examinations and 67 on the Nelson-Denny. These points, based upon -1 standard deviation in the local population, are somewhat higher than the national experience. Direct comparisons to the national data are not made, though, since the tests are given nationally for so many reasons that such comparisons are not justified.

The data utilized in the study was transcript information on 1195 Kansas public university students who had attempted the Nelson-Denny Reading Test and/or the Pre-Professional Skills Test(s) during the 1983-85 academic years. Table 1 lists the demographic distribution of the

COMPETENCY TESTING
population.

Table 1

Characteristics of the Student Population

Unit	No. Subjects	Mean GPA	% Pass		
			N-D	PPST WRIT	PPST MATH
LAS	318	2.91	82.1	80.44	83.5
Education	405	2.99	78.8	86.2	85.8
Business	434	2.90	73.4	84.3	84.9
Unclass.	38	2.85	70.1	78.1	94.4

To study the phenomenon of success/failure, ACT scores, competency test results, and general education enrollment patterns were extracted from each transcript. In keeping with the exploratory nature of the study, selected general education core course grades were coded, as were personal data such as student classification, GPA, and major.

General education courses taken in each of four basic areas--humanities, social science, language arts, and natural sciences--were coded by credit hours in the field, and student classification when the work was completed. This data was transformed into a measure of curricular breadth for each student. This was accomplished by calculating the proportion of general education hours taken in each area and obtaining a coefficient of variation for each case. The depth of the general education experience was defined as the total number of non-major freshman-sophomore courses completed. Timeliness of enrollment was the proportion

COMPETENCY TESTING

of these hours the individual student had completed by the end of the sophomore year.

A subset of 87 students was randomly chosen from the population, and the Text Anxiety Scale was administered to them. Results of this testing were added to the existing data.

Statistical Design

Descriptive statistics, means, standard deviations, and frequencies were used to describe the data base. In addition, these measures were used to describe the historical experience of the student body in regard to the variables selected for study. Simple correlation matrices were established to describe variable relationships. A multiple correlation analysis was performed to assess both the strength and the relative importance of the discrete variables as predictors of competency test scores. The efficiency of each equation was assessed by completing a series of discriminant analyses regressing the identified variables on subject success/failure with the three competency exams. The criterion in this instance was the proportion of correct predictions of observed test outcomes.

The extremely high correlations obtained between ACT scores and the various standardized measures taken tended to attenuate the other variables in the data base. Therefore, a similar analysis was completed utilizing demographic and general education information, but without the inclusion of ACT data.

The records of international students were retrieved

COMPETENCY TESTING

from the data base and studied as a group to determine whether the parameters describing their experience differed from whether the parameters describing their experience differed from those of native speakers. Initial intent was to apply univariate tests of significance to this data, but at the descriptive level such pronounced differences were encountered that meaningfulness was established without application of such techniques.

Results

The statistical analysis was conducted in several dimensions. Quartiles of ACT composite distribution were cross tabulated with PPST results:

Table 2

ACT Quartiles Compared with PPST Writing Status*

	ACT Range	PPST Pass	PPST Fail
Q1	8-15	63%	27%
Q2	16-18	90%	10%
Q3	19-21	95%	5%
Q4	22-31	99%	1%
Total Pass-Fail Rate		89%	11%
Mean PPST Scores		176 S=6.1	166 S=12.5

COMPETENCY TESTING

Table 3

ACT Quartiles Compared with PPST Math Status*

	ACT Range	PPST Pass	PPST Fail
Q1	8-15	36%	64%
Q2	16-18	85%	15%
Q3	19-21	99%	1%
Q4	22-31	97%	3%

Total Pass-Fail Rate

Mean PPST Scores	180 S=8.2	166 S=3.5
------------------	--------------	--------------

Table 4

ACT Quartiles Compared with Nelson-Denny* Pass-Fail Rate

	ACT Range	N-Denny Pass	N-Denny Fail
Q1	8-15	52%	48%
Q2	16-18	79%	21%
Q3	19-21	96%	4%
Q4	22-31	100%	0

Total Pass-Fail Rate 82% 18%

Nelson-Denny Mean Scores	95 s=20.0	56 s=8.5
--------------------------	--------------	-------------

*Percentages vary from total reported in Table 1, as not all students report ACT scores.

Several observations can be made about these tables. It can be noted that the Nelson-Denny test appears to act most efficiently of the three tests as a scholastic aptitude measure. There is a wide spread between the mean scores achieved by successful and unsuccessful students, and fairly low variability in the scores of those who failed. The pattern of pass vs. fail is quite consistent and dramatically discriminatory as scholastic aptitude increases.

COMPETENCY TESTING

The Pre-Professional Skills Tests, however, appear to have somewhat less power to discriminate scholastic ability. Examination of the means and standard deviations indicate considerably more distribution overlap. Indeed, on the math test a higher proportion of Q3 students achieved a pass status than was the case with Q4 students. Considering the experience of students by ACT composite score increments gives credence to this interpretation, as there is considerable variation in the pass/failure rate as the ACT standard score distribution is ascended. For example, for students sitting for the English examination who had attained ACT scores of 19 (3rd quartile) the proportion, as well as the number, of failures was less than for those with ACT scores of 20 or 21. This same pattern existed across the first three quartiles of both the math and the English exam.

Based upon these three tables, several generalizations can be drawn. First, it appears that the tests are measuring something in addition to scholastic aptitude, or to state this point in another direction, raw scholastic aptitude is not a completely accurate indicator of measured competency except at the extremes of the ability spectrum. Therefore, if it is accepted that the tests do measure that quality of performance known as competency, and if the pass-fail cut scores are valid, the tests cannot be replaced with simple scholastic aptitude measures for the majority of the student body. On the other hand, if competency is, as Weaver (1980) contends, a function of public perception rather than an

COMPETENCY TESTING

intellectual constant, it is possible that what is appropriate to measure is an intelligence related factor such as scholastic ability.

Multivariate and Univariate Analysis of Contributing Factors

To provide a starting point for multivariate analysis of the data, zero order correlation coefficients were run between all variables collected. Table 5 is the resulting matrix.

Table 5
Inter Item Correlations

	EN 1	EN 2	MA 1	PPST GPA	PPST Eng.	Nelson Math	ACT Denny	ACT Eng.
EN 1	1.0							
EN 2	.488	1.0						
MA 1	.275	.358	1.0					
GPA	.553	.582	.506	1.0				
PPST	.342	.221	.159	.167	1.0			
English								
PPST	.260	.139	.426	.281	.212	1.0		
Math								
Nelson	.351	.379	.224	.419	.249	.336	1.0	
Denny								
ACT	.435	.348	.285	.425	.409	.518	.625	1.0
English								
ACT	.291	.264	.427	.398	.272	.693	.474	.53
Math								

The most noticeable feature of the matrix is the comparative weakness of the individual coefficients, as contrasted to the extremely consistent directionality. It is suggested that all measures are related to the same general area, but are measuring those areas (e.g., scholastic aptitude) from a different perspective. One interesting, and rather curious, statistic is the correlation between the ACT English scores

COMPETENCY TESTING

and the PPST writing score, as opposed to the relationship between ACT and PPST math scores. In the latter case the measures covary at a 50% or very high level, while in the former the relationship is 16%. This indicates that while both tests are measuring approximately the same thing in the same way in math, they are measuring something quite different in the English area.

Multiple regression analysis was performed on each competency examination by regressing items in the data base on each set of exam scores to determine if it was possible to build a predictive equation to assess early a student's propensity to succeed or fail on the examinations. Since the emphasis was upon prediction, all variables which appeared to provide a greater than chance discriminatory power were retained for study.

In the case of the PPST Math examination, the variable with the greatest predictive power was the ACT math score, accounting for some 45% of the variance by itself. The next best variable indicated by the analysis was GPA, followed by Math 1 grade. However, further analysis indicated that GPA was more likely acting as a suppressor on the math grade variable than as a true predictor. This observation was confirmed later in the analysis when the sufficiency of the individual indicators was assessed.

In investigating the PPST English examination, the analysis indicated that the predictor variables could account

COMPETENCY TESTING

for approximately one-third of the variance. Once again, the best predictor was the ACT English score, with the second most powerful being the English 1 grade. No other variables contributed sufficient stability to the equation to be considered.

The analysis of the Nelson-Denny examination revealed that two ACT scores, Social Studies and English, were the most powerful predictors, followed by English 1 grade. By far the most powerful was ACT social studies accounting for approximately two-thirds of the explained variance.

While the above analyses do explain the relationship of the various factors in the design, the sufficiency of the combinations of variables as actual working models to predict success or failure was not clearly assessed by the technique. To make this determination, a discriminant analysis was performed on the data to ascertain how well each set predicted the actual experience of the subjects. This type format is commonly known as a percentage of correct placements analysis. Tables 6, 7, and 8 indicate the results.

Table 6

Percentage of Correct Placements from
Pre-Determined Variable Set, PPST English Exam*

	Correct Prediction	Incorrect Prediction
Pass	99.1%	.9%
Fail	88.9%	11%

*ACT English Score and English 1 Grade

COMPETENCY TESTING

Table 7

Percentage of Correct Placements from Pre-Determined Variable Set, PPST Math Exam*		
	Correct Prediction	Incorrect Prediction
Pass	99.3	6.7%
Fail	66.7%	33.3%
*ACT Math and Math 1 Grade		

Table 8

Percentage of Correct Predictions from Pre-Determined Variable Set, Nelson-Denny Reading Exam*		
	Correct Prediction	Incorrect Prediction
Pass	94.8%	5.2%
Fail	55.9%	44.1%
*ACT Social Studies, ACT English and English 1 Grade		

Three features should be noted from these tables. Most importantly, when combined with the examination of the frequency distributions it must be observed that the strength of the ACT predictors will be differential across the distributions, being most efficient at the two extremes. Additionally, in all cases the equations predict success much more strongly than they predict failure. This would tend to suggest that something other than scholastic aptitude (ACT scores) or performance-motivation (class grades) is operational, and that this something is a much more significant factor in the case of those who fail the PPST Math and the Nelson-Denny test, than the PPST English test. However, a third factor which should be recognized is that with all three tests, but particularly with the PPST instruments, the predictive power achieved through knowledge of the student's ACT scores and course grades is considerably greater than chance.

COMPETENCY TESTING

Enrollment Patterns as Predictor Variables

General education enrollment patterns were analyzed to determine whether such patterns affected success on the competency examinations. As might be expected, the most predictive variable for any given test was the total number of credit hours taken in the area being measured by the examination. Thus, the best predictor of score on the PPST Writing examination was total hours of language credit, while hours in math/science was the best predictor of PPST Math score. Interestingly, as noted in other studies of a similar nature (McPhee & Kerr, 1985), social studies training serves as a secondary predictor of math proficiency.

Discriminant analysis indicated extreme efficiency of a predictive equation using proportional distribution of general education courses and timeliness of enrollment as predictors. Successful students on all three examinations showed a nearly equal distribution of natural sciences/math; language/humanities; and social sciences in their general education program. Unsuccessful students tended to have enrollment patterns which upset this balance, and had a tendency to specialize in one area during their freshman-sophomore years at the expense of the other two.

Grade Point Average as a Predictor

Another area of interest was the predictive power of GPA. Since a GPA of 2.5 is necessary for graduation, a case can be made for this variable being the true measure of student competence. Tables 9, 10, and 11 report the results

• COMPETENCY TESTING

of a comparison of these measures.

Table 9

PPST Math Status by GPA Category

GPA	Pass	Fail
Above 2.5	90%	10%
Below 2.5	72%	28%

Table 10

PPST English Status by GPA Category

GPA	Pass	Fail
Above 2.5	92%	8%
Below 2.5	69%	31%

Table 11

Nelson-Denny Status by GPA Category

GPA	Pass	Fail
Above 2.5	84%	16%
Below 2.5	66%	34%

It is interesting to note that while GPA does tend to discriminate, the pattern of the discriminant analysis--results--much higher discrimination of success, but considerably lower of failure--still tends to hold true. A peculiarity in the data is the apparent stronger discriminatory power of GPA alone as opposed to the discriminant equation in predicting the unsuccessful condition. This is probably attributable to missing data in the discriminant analysis rather than power. Careful examination of the two methods of predicting success/failure indicates that the discriminant equations are much stronger overall predictors than is GPA in forecasting passing/failure.

• COMPETENCY TESTING

Another area of interest is test anxiety. Carrier and Jewell (1966) found that anxiety is related to academic performance, and in the case of academic tests success can be predicted by self reports of anxiety. If this is indeed the case, it seems logical to assume that if the tests are measuring competency, the area other than scholastic aptitude being measured may be the individual's feelings of security, or lack of academic anxiety. Current research supports this position (Zigler & Trickett, 1978). This area was studied by selecting a random sample of 87 students and administering the Test Attitude Inventory Table 12 illustrates correlations between test anxiety and the various competency measures.

Table 12

Correlations Between Competency Test Scores
and Test Anxiety Scale Scores for
Random Sample of Test Takers
(n=87)

	PPST Writ.	PPST Math	N-D	Total	Worry
PPST Writ.	1.0				
PPST Math	.21	1.0			
Nelson-Denny	.25	.34	1.0		
TAS Total	-.41	-.48	-.50	1.0	
TAS Worry	-.46	-.48	-.47	.91	1.0
TAS Emot	-.36	-.48	-.45	.96	.81

The Test Anxiety Scale Total Score is indicative of overall anxiety, which can be thought of as insecurity, fear, and/or dread of the testing process. Emotionality is a subscore which focuses particularly upon the testing situation, while Worry is a subscore which assesses insecurity of outcomes. Research has demonstrated that these measures are

* COMPETENCY TESTING

predictive of academic achievement, and has shown a relationship between such measures and intelligence. However, these studies have all been conducted using course grades as a criterion. The experience with this group indicates a much stronger relationship between text anxiety and competency test scores than has been previously assumed.

Since this data set is too small to subject to multivariate analysis, mean analysis was used to compare differences between those who passed and those who failed. Tables 13, 14, and 15 indicate results of this comparison. It should be noted that a higher score indicates greater anxiety.

Table 13

Comparison of Test Anxiety Scale Scores
for Students Passing and Failing the
Nelson-Denny Reading Examination

	Mean, Pass	Mean, Fail	t	p
Total	37.0	49.4	-3.5	.005
Worry	13.4	17.6	-3.1	.005
Emotionality	15.5	21.3	-3.9	.005

Table 14

Comparison of Test Anxiety Scale Scores
for Students Passing and Failing the
PPST Math Examination

	Mean, Pass	Mean, Fail	t	p
Total	37.9	50.8	-8.4	.005
Worry	13.5	18.6	-4.1	.005
Emotionality	16.0	21.6	-3.7	.005

• COMPETENCY TESTING

Table 15

Comparison of Test Anxiety Scale Scores for Students Passing and Failing the PPST Writing Examination

	Mean, Pass	Mean, Fail	t	p
Total	39.0	48.8	-3.0	.005
Worry	13.7	18.9	-3.5	.005
Emotionality	16.7	20.0	-1.9	.05

While there is considerable debate as to just what these test anxiety scores mean and how they should be interpreted, comparison of the sample demographics with those reported in the literature indicate that it is extremely unlikely that the reported values are merely a function of scholastic aptitude. In fact, it may well be possible that test anxiety is responsible in large part for the observed differences in performance of students of like scholastic ability.

The Situation Regarding International Students

International students present an unusual problem in regard to the competency examinations. The data base contains information on 35 such individuals, and only one of those has completed the ACT examination. Therefore, little can be done with this data other than describing it. Of the students represented, 20 are Arts and Science, 13 are Business, and 2 are unclassified students. As a group they have achieved a mean GPA of 2.93, with 14 having GPA's exceeding 3.0. All but two of the group completed EN 1. The highest grade achieved in that course was an "A" (f=6) and the lowest a "D" (f=13), with a mean grade of 2.79. Twenty-

• COMPETENCY TESTING

eight of the group continued through EN 2, and performed somewhat better with two earning "A's", but only 6 receiving "D's", and the mean grade increasing to 2.86. Thirty-two of these students had attempted the PPST writing examination, with only five (15%) being successful.

The results on the Nelson-Denny Reading Test were even more startling. Of the 15 who attempted this examination, only one passed. It should be noted that no data was collected on the culture-free alternative, the Michigan Test.

International students fare considerably better in the field of mathematics. The mean grade earned by the 17 who completed MA 1 was 2.94, with a median grade of "B" achieved. Of the 17 who attempted the PPST Math Test, 12 passed.

The data concerning international students is very-s sketchy, and the total number of subjects is so small that complex analysis is not possible. Table 16 demonstrates these differences in experience of foreign students and native speakers.

Table 16

Comparison of Experience of Native Speakers
and International Students

	Native Speakers	International Students
Combined GPA (Mean)	2.60	2.93
EN 1 Grade (Mean)	2.61	2.79
EN 2 Grade (Mean)	2.79	2.86
PPST Writing Pass	83.8%	15.1%
Nelson-Denny Pass	78.8%	6.6%
MA 1 Grade (Mean)	2.61	2.94
PPST Math Pass	84.9%	70.6%

Study of these data lead to disturbing conclusions. The

COMPETENCY TESTING

GPA represents some 30 separate, and largely independent, measures on scholastic ability which are gathered from across the campus. In this regard, it appears that the overall demonstrated proficiency of international students is slightly higher than that of native speakers. Additionally, if the problem is nothing more than a simple inability to manipulate the language, this fact is not being reflected in English courses, where, again, the performance of international students exceeds that of native speakers. This superior performance is even more striking in the field of mathematics. However, the success of international students is poor beyond reasonable explanation on the competency examinations. This can only lead to the conclusion that the tests or the testing situation itself, contains a cultural bias. The use of the instruments to assess competency with international students is therefore questionable.

Discussion

Data analyses allow conclusions to be drawn regarding the four major questions posed at the beginning of this article. First, competency tests appear to primarily differentiate students on the basis of scholastic ability. The categorization of students is sharpest with the Nelson-Denny examination and somewhat less with the Pre-Professional Skills Test. While it appears that the PPST Math Test results are more understandable in terms of scholastic aptitude than the English examination, both tests seem to work

* COMPETENCY TESTING

most effectively at the two extremes of the scholastic aptitude continuum. At the upper and lower end of the ability scale, the information yielded by the competency exams is redundant with similar information yielded by ACT scores. Or, to state the situation in different terms, there seems to be a threshold level at both ends of the ability continuum where the tests tend to become purely scholastic aptitude measures. The range of this zone at the upper end of the continuum is far greater than is the case with the lower end.

Since the tests work differently at the extremes of the distribution, a strong possibility exists that patterns of relationships within the central part of the data set are attenuated in this study. For example, while it appears that scholastic ability accounts for approximately half of the success or failure on competency examinations, it may be that most of this can be attributed to the experience of bright or slow students, and for those between the extremes something other than scholastic ability is the primary factor. As funds and time permit, it would seem reasonable to isolate the students in the central part of the distribution and study their characteristics separately.

A second conclusion which can be drawn is that the multiplicity of tests administered provides little "new" information about students. Utilizing cut scores on entrance examinations, it would be possible to excuse a considerable number of students from the competency testing requirements, thus saving considerable time and money for all concerned.

COMPETENCY TESTING

A third conclusion is that variables do exist which if evaluated early in the student's academic career can increase their success on competency tests. Extremely low entrance examination scores can be used to assign students to high risk categories. Treatment tailored to the specific nature of the individual's needs can be applied in advance of the time during the student's career when competency tests are given. Passing scores appear to be related to enrollment in courses relevant to the type of test taken. Language arts courses correlate highly with the PPST Writing exam, humanities coursework correlates with the Nelson-Denny Reading exam, and natural science courses correlate with the PPST Math exam. The University of Arkansas at Pine Bluff has had significant success in improving test scores of similar students through teaching test taking strategies (Hackley, 1985). This would seem to fit extremely well with Astin's (1982) value-added model of excellence in education. Further information should be gathered concerning the nature of student characteristics which impact on successful/unsuccessful experience on competency examinations. Of particular value would seem to be studies of test anxiety and patterns of general education enrollment. In general, breadth of general education preparation tends to improve competency test performance. It is recommended that the competency testing requirement not be applied to international students, but rather be offered as an option. While it is not possible

• COMPETENCY TESTING

at this time to define the problem, it is obvious that all three examinations seem to contain systematic bias when applied to international students. If, indeed, there is institutional concern regarding proficiency of foreign students, the answer would seem to lie in raising admission requirements. Clearly, as the data supports, the use of an exit exam is not a viable alternative.

A most disturbing extension of these conclusions is that the study indicates that what is being measured by the new "competency" examinations is not any new dimension of student behavior, but rather the same type of ability-achievement-aptitude which is measured in so many ways in education. However, if as this study suggests, students who are otherwise in good standing are being eliminated by tests designed to measure minimal scholastic aptitude, the entire system of entrance and quality control standards currently in place is open to question. Since this cutting of students occurs fairly late in the students' college career, it is imperative that further work be conducted to define competency, and to ascertain the most effective way to measure the concept.

Unless this is done it appears that for the students in the center of the ability spectrum, what is being measured by the standardized tests may not relate to competence, while for the students at the extremes, the examinations are useless hurdles to be overcome.

Dealing with a competency testing program is problem-

• COMPETENCY TESTING

atic. Obviously, the ideal situation would be one in which extremely rigorous instruments are used, but all students pass. The reality of both human nature and the frailty of empirical measures is such that a system of this type cannot exist. But, it seems, all institutions strive toward such a Utopian design.

References

- Astin, Alexander W. (1982, November). Excellence and equity in American education. Paper presented at a meeting of the National Commission on Excellence in Education. (ERIC Document Reproduction Service No. ED 227 098)
- Bestor, Arthur (1974). In defense of intellectual integrity. In Hook, Sidney; Kurtz, Paul; & Todorovich, Miro (Eds.) The Idea of a Modern University. Buffalo, NY: Prometheus Books.
- Biemiller, Lawrence (1982). The unmaking of a president, 1982: Tumult at West Chester State. The Chronicle of Higher Education. Sept. 29, 4-6.
- Carrier, Neil & Jewell, Donald (1966). Efficiency in measuring the effect of anxiety upon academic performance. Journal of Educational Psychology. 57, 23-26.
- Cashen, Carol J., Fillipone, Samuel R., & Gajewski, Geoffrey C. (1981). From basic to collegiate skills. Journal of Developmental and Remedial Education. 4 (2), 22-24.
- Cross, K. Patricia (1971). Beyond the open door. San Francisco: Jossey-Bass.
- Hackley, Floyd (1985, August). Quality education and the preparation of teachers. Paper presented at the Summer Workshop of the Association of Teacher Educators, Wichita, KS. (Author)
- Jacobsen, Robert L. (1982). Massachusetts system thrown into a turmoil over finances and governance. The Chronicle of Higher Education. February 10, 4-5.
- Kraetsch, Gayla A. (1980). What's new in literacy legislation? Change. April, 44-45.
- Magarrel, Jack (1982). More colleges marshal data to demonstrate their economic impact. The Chronicle of Higher Education. March 10, 8.
- Machlup, Fritz (1974). The illusion of universal higher education. In Hook, Sidney; Kurtz, Paul; and Todorovich, Miro (Eds.) The Idea of a Modern University. Buffalo, NY: Prometheus Books.

- McPhee, Sidney & Kerr, Michael E. (1985). Scholastic aptitude and achievement as predictors of performance on competency tests. Journal of Educational Research. 78, 186-190.
- Mingle, James R. ed.(1981). Challenges of retrenchment. San Francisco: Jossey-Bass.
- Moos, Malcom (1981). The post land grant university: The University of Maryland Report. Adelphi, MD: Office of University Publications.
- National Commission on Excellence in Education (1983). A Nation at risk: The imperative for educational reform. (No. 065-000-00177-2) Washington, D.C.: U.S. Government Printing Office.
- Preer, Jean L. (1983). Competency, admissions and articulation: Returning to the basics in higher education. (ASHE-ERIC Research Report No. 6) Washington, D.C.: Association for the Study of Higher Education.
- Watkins, Beverly C. (1983). Mastery of six basic subjects and six intellectual skills urged for college-bound seniors. Chronicle of Higher Education. May 18.
- Zigler, E. & Trickett, P.K. (1978). I.Q., social competence and evaluation of intervention programs. American Psychologist. 33, 789-798.